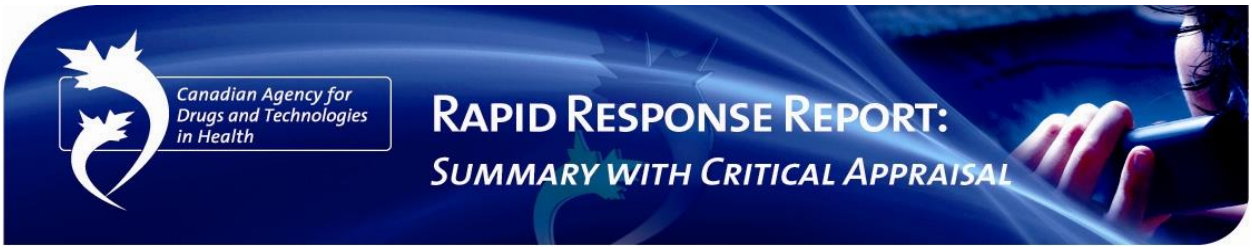


TITLE: Diagnostic Methods for Neuropathic Pain: A Review of Diagnostic Accuracy

DATE: 07 April 2015

CONTEXT AND POLICY ISSUES

Compared to nociceptive or inflammatory pain, individuals with neuropathic pain (NP) suffer from more severe disease, greater costs, and relatively reduced health related quality of life.[1] Direct and indirect costs of NP represent a substantial economic burden on the Canadian healthcare system with per patient costs estimated at $2567 (± $2711) per three month care period.[2] Neuropathic pain is defined by the International Association for the Study of Pain (IASP) as "pain arising as a direct consequence of a lesion or disease affecting the somatosensory system".[3] The etiology of NP is broad and associated conditions (e.g., cancer, surgery, diabetes, herpes zoster)[4] have been classified into four distinct categories: peripheral nervous system focal and multifocal legions (e.g., post-herpetic neuralgia); peripheral nervous system generalized polyneuropathies (e.g., diabetic neuropathy); central nervous system lesions (e.g., spinal cord injury); and, complex neuropathic disorders including complex regional pain syndrome types I and II.[5,6] While rates of NP-associated conditions are well documented, rates of NP in the general population are difficult to quantify and under-diagnosed.[7] Studies in the United Kingdom and France that utilized screening tools to identify NP have estimated that 6-8% of patients with chronic pain experience NP in the general population.[8,9] A single Canadian study that used telephone-based questionnaires for determining NP rates estimated a higher (18%) rate in the general population.[10] Limitations and lack of standardization of diagnostic methods increase the potential for undetected or poorly classified cases. There is no recognized objective gold standard for assessing NP.[11] However, the Special Interest Group on Neuropathic Pain (NeuPSIG) of the IASP has set out a grading system that has been used to guide clinical assessment and diagnosis.[3,12] This approach involves multiple steps including obtaining a clinical history of pain, assessing the neuroanatomical plausibility of pain, using sensory assessments to confirm nervous system involvement, and running diagnostic tests to confirm nervous system lesions or disease. Other less resource intensive methods of diagnosis have been documented and may be especially useful in primary care.[13] These strategies include, but are not limited to NP screening tools such as: the Douleur Neuropathique 4 (DN4),[14] PainDETECT (PD-Q)[15] the Leeds Assessment of Neuropathic Symptoms and Signs (LANSS),[16] and the standardized evaluation of pain (StEP).[17] Screening tools are comprised of an interview component and, in some cases, the addition of a brief bedside clinical assessment. Many of

these tools have been translated for application in other languages and populations.[18] Given observed variation in the approach to diagnosis of NP and the significant disease burden, it is of interest to assess the diagnostic accuracy of methods of assessing NP.[19] Improved diagnostic procedures may facilitate improvements in treatment approaches.

RESEARCH QUESTION

What is the accuracy of diagnostic methods for neuropathic pain in adults?

KEY FINDINGS

Fourteen studies were identified regarding the accuracy of diagnostic methods for neuropathic pain in adults. Accuracy of diagnostic methods varied substantially by clinical population and tool. Overall, there is evidence to suggest superiority of the DN4 screening tool. Diagnostic screening tools, particularly the LANSS and painDETECT, were not deemed sufficiently accurate for identifying NP in cancer, fibromyalgia, failed back surgery syndrome, and upper limb and back pain. There is no consensus on a universally appropriate screening tool for NP. Risk of bias, particularly regarding the administration of the index and reference tests, should be taken into account in interpretation of the results.

METHODS

Literature Search Methods

A limited literature search was conducted on key resources including PubMed, The Cochrane Library (2015, Issue 3), University of York Centre for Reviews and Dissemination (CRD) databases, ECRI, Canadian and major international health technology agencies, as well as a focused Internet search. Methodological filters were applied to limit retrieval to health technology assessments, systematic reviews, meta-analyses, randomized controlled trials (RCTs), and non-randomized studies. Where possible, retrieval was limited to the human population. The search was also limited to English language documents published between January 1, 2010 and March 2, 2015.

Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

Table 1: Selection Criteria

Population	Adult population with neuropathic pain (NP) symptoms
Intervention	Diagnostic tests for NP (including but not limited to: reflex testing, superficial pain testing, light touch perception, vibration testing, sympathetic skin response, quantitative sensory testing, nerve conduction studies, testing of the autonomic system, electrophysiologic studies, NP screening tools)
Comparator	Standard reference tests (e.g., diagnosis by clinician, International Association for the Study of Pain clinical grading system)
Outcomes	Diagnostic accuracy test measurements (i.e., sensitivity, specificity, positive

	predictive value [PPV], negative predictive value [NPV], likelihood ratios [LR], area under the receiver operating characteristic curve [AUROC], Youden's index, diagnostic odds ratio [DOR])
Study Designs	Health technology assessment reports, systematic reviews, meta-analyses, randomized controlled trials, non-randomized studies

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, they were duplicate publications, there was no reference test, or were published prior to 2010. In addition, studies were excluded if they were conducted to adapt or assess the validity or diagnostic accuracy of a translated screening tool. These translation studies are listed in Appendix 5. Conversely, all studies that assessed diagnostic accuracy in specific clinical populations were included. Studies that assessed the accuracy of diagnostic tools for factors associated with NP such as peripheral neuropathy were excluded.

Critical Appraisal of Individual Studies

The included diagnostic studies were critically appraised using the QUADAS-2 checklist.[20] Summary scores were not calculated for the included studies; rather, a review of the strengths and limitations of each included study were described and presented in tabular and graphical format.

SUMMARY OF EVIDENCE

Details of study characteristics, critical appraisal, and study findings are located in Appendices 2, 3, and 4, respectively.

Quantity of Research Available

A total of 720 citations were identified in the literature search. Following screening of titles and abstracts, 685 citations were excluded and 35 potentially relevant reports from the electronic search were retrieved for full-text review. Three potentially relevant publications were retrieved from the grey literature search. Of these 38 potentially relevant articles, 24 publications were excluded for various reasons, while 14 publications met the inclusion criteria and were included in this report. Appendix 1 describes the PRISMA flowchart of the study selection.

Additional references of potential interest are provided in Appendix 5.

Summary of Study Characteristics

Appendix 2 provides details on the study characteristics.

Fourteen diagnostic accuracy studies were identified regarding the accuracy of diagnostic methods for NP in adults. Cross-sectional assessments were done for all diagnostic accuracy evaluations. Several studies were embedded within larger trials, or were conducted as follow up analyses to RCTs.

Country of Origin

All studies, excluding three[21-23] were conducted outside of North America.

Study Setting

Studies were conducted in a range of settings including primary care,[24] hospitals[21,25,26] and other secondary care settings,[22,27-29] field settings,[30] and tertiary care (e.g., injury and rehab centers).[6,23,31-33]

Patient Population

Studies included a range of adult (≥18 years) clinical populations with suspected neuropathic pain. This included general chronic pain populations,[24,27,32] patients with condition related pain (including leprosy, cancer, fibromyalgia, and diabetes),[6,21,25,28-31] and patients with injury or failed surgery related pain.[22,23,26,33]

Interventions and Comparators

The majority of included studies assessed neuropathic pain screening tools including the DN4,[21,22,24,25,27,29-31,33] LANSS,[22,25,27,31-33] and PD-Q.[6,28,31-33] Most studies assessed more than one screening tool. In addition, one study assessed a spinal cord injury specific screening tool.[23] One study assessed an abbreviated 4 step-tool based on IASP clinical criteria.[24] One study assessed the diagnostic accuracy of a groupings of symptoms and signs of NP.[26] Index tests were universally compared against various clinical diagnostic methods as a reference standard. These clinical assessments included IASP grading criteria,[21,23,24,31,32] the Quebec Task Force Classification of Spinal Disorders,[22] the European Federation of Neurological Societies guidelines,[28,33] and the Edmonton Classification System for Cancer Pain.[6] Some studies used other less specific processes involving clinical opinion.[25-27,29,30] For example, several studies simply described the reference test as being based on expert clinical opinion without describing the exact diagnostic process, whereas others outlined a specific set of diagnostic procedures without attributing specific guidelines.

Outcome Measures

All studies assessed sensitivity and specificity and many included additional measures of diagnostic accuracy including positive and negative predictive value, area under the receiver operating characteristic curve, Youden's index, and diagnostic odds ratios.

Summary of Critical Appraisal

Appendix 3 provides details on critical appraisal. QUADAS-2[20] assessment results are presented in tabular (Table A2) and graphical (Figure A1) format.

There were minimal issues with patient selection. Some studies intentionally excluded hard to diagnose patients[23,26] and patients with advanced disease.[6] This may have led to overestimation of diagnostic accuracy. Also, some studies did not consecutively recruit patients, using ad hoc methods and increasing the risk of selection bias.[26,31] In terms of applicability, some studies may not be representative total clinical populations. For example, clinical populations with possible

mixed pain,[23,26] patients with advanced disease,[6] and patients with a history of NP[28] were excluded from some studies.

Few studies were without risk of bias stemming from conduct of the index test. Two studies used the same assessor or clinician to perform the reference and index tests.[21,26] One study failed to blind the index assessor to the results of the reference test[32] and for several studies blinding of assessors was unclear.[6,25,28,30] Given the baseline subjectivity of these screening tools, lack of blinding may have opened these studies up to detection bias. Two studies used multiple assessors for the index assessment. In these cases it is possible that issues with inter-rater reliability may have skewed the results.[23,27] In addition, thresholds were optimized in several instances which may have resulted in enhanced estimates of diagnostic accuracy. In general, the index tests applied were appropriate for the review question though it was not possible to rule out potential differences in administration of the screening tool and the impact on diagnostic accuracy.

There were major concerns with the administration of the reference standard for most studies. In two studies the credentials of the assessor was unclear.[21,30] Clinical diagnosis requires significant expertise and it may be inappropriate for untrained individuals to attempt this process. One study used different clinicians to perform the reference test without reporting inter-rater reliability.[23] One study used multiple assessors collaboratively, which may not accurately reflect typical practice.[33] One study reported varied application of the reference test with some patients undergoing diagnostic tests while others received an incomplete assessment, resulting in more conclusive diagnosis for a subset of patients.[31] Several studies failed to blind the reference assessor to results of the index test or failed disclose blinding strategy.[6,25,26,28,30] Several studies did not use the most up to date guidelines for clinical diagnosis of NP.[27,29,33] In some cases this was due to lack of availability of these guidelines at the time of publication. The appropriateness of the reference test in these studies is unclear. Subsequent diagnostic accuracy studies may benefit from use of the updated reference test guidelines. In terms of applicability, several studies either failed to properly describe their reference assessment process or failed to provide evidence of their appropriateness for NP.[26,30]

In terms of flow and timing, several studies failed to report the order of the reference and index tests or what the duration between tests was.[22,23,25-28] In situations where blinding was unclear this makes it difficult to discern whether the assessment of the index or reference test was compromised. Also, this introduces the possibility of change in condition over the time between tests.

Summary of Findings

Appendix 4 provides detailed outcome data and author's conclusions.

What is the accuracy of diagnostic methods for NP in adults?

Diagnostic accuracy of the assessed screening tools and classification approaches varied significantly by strategy and by population. Overall, the DN4 was the most universally well performing screening tool, showing reasonable accuracy (sensitivity: 76 to 100%; specificity: 45 to 92%) in all patient populations in which it was assessed except those with failed back surgery syndrome (sensitivity: 62%; specificity: 44%).

Condition-Related Pain

Two studies reported good sensitivity (85 to 100%) of both the LANSS and DN4 for detecting NP in patients with leprosy.[25,30] However, this was at the expense of specificity (42 to 45%) in one of the studies.[25] In cancer patients, the DN4 performed well (sensitivity: 88%; specificity: 88%).[31] According to the authors of the reports, the specificities of the LANSS and painDETECT were insufficient (18 to 68%) to recommend use in this population.[6,31] In patients undergoing breast resection surgery the DN4 had good sensitivity (90%) but only moderate (60%) specificity.[21] In fibromyalgia patients, it was reported that the PD-Q screening tool did not have good diagnostic capacity (sensitivity = 79%; specificity = 53%).[28] In patients with diabetes the DN4 full and interview-only screening tools performed well (sensitivity and specificity >80%).[29]

Injury or Surgery-Related Pain

Overall, the performance of the DN4 was superior to other screening tools (LANSS and the Neuropathic Pain Questionnaire) in patients with spinal cord injury. The DN4 showed good sensitivity (88 to 93%) and reasonable specificity (75 to 88%) when applied in spinal cord injury patients.[23,33] The AUROC values suggested that on average spinal cord injury patients with NP will have more abnormal screening results than 77 to 86% of those without NP. [23,33] One study reported accuracy values suggestive of poor performance of the LANSS, PD-Q, and the Neuropathic Pain Questionnaire, particularly in terms of sensitivity (8 to 50%).[33] In patients with failed back surgery syndrome the DN4, LANSS, and both screening tools combined showed poor to moderate accuracy measures.[22]

General Non-Attributed Pain

In patients with chronic pain the IASP grading system-derived four question screening tool showed poor sensitivity (52%) and reasonable specificity (79%).[24] The DN4 performed better in this patient population with reasonable sensitivity (76 to 87%) and moderate specificity (57 to 70%).[24,27] The LANSS showed good specificity (94%) with only moderate sensitivity (76%) suggesting that the DN4 may be preferred.[27] In patients with upper limb and neck pain neither the LANSS nor PD-Q questionnaires performed well.[32] For patients with low back or leg pain, a model derived cluster of signs and symptoms performed well with good sensitivity and high specificity. For this test, the odds of a positive NP rating were 150 times higher in those with NP than those without.[26]

Limitations

Due to the scope of the report, which focused on neuropathic pain, studies assessing the diagnostic accuracy of specific neurophysiologic or electrophysiologic tests or tools used for assessing components of NP diagnosis - such as nervous system dysfunction or neuropathy - were not captured. For example, studies assessing the diagnostic accuracy of neurophysiologic tests to determine the presence of nerve lesions or nervous system dysfunction but not the overall presence of neuropathic pain have not been reviewed.

Based on the risk of bias and applicability assessments, there were minimal issues with the patient selection in most studies, although pre-selection of patients with probable NP or with less complex disease may have resulted in overoptimistic diagnostic accuracy results in some instances. There were substantial issues with the conduct of the index and reference assessments in some studies. Approximately one third of the included studies failed to blind

assessors or disclose blinding procedures, suggesting a potential for detection bias. The qualifications of the reference test assessor were not always clearly stated and the most up to date guidelines for NP diagnosis were not always used. Proper clinical diagnosis requires the judgement of a pain specialist and all components of the clinical diagnostic process. In the same vein, several instances of multiple assessors may have introduced inconsistency. There was also some uncertainty regarding the flow and timing of the patients and assessments. Where these factors were unclear it was not possible to rule out changes in the patient`s condition between tests or influence of the results of the earlier assessment on the final judgement for the subsequent test.

Lastly, the generalizability of the condition specific findings presented in this review should be acknowledged. Comparative diagnostic accuracy of these tools in different conditions was not assessed; therefore, no conclusions could be made about their relative performance across clinical populations. Evidence of diagnostic accuracy in one condition may not translate to other conditions as is evident from the variability in the results presented. Similarly, not all screening tools were assessed in all conditions; therefore, accuracy of one tool in a clinical population does not ensure that other tools will perform well.

CONCLUSIONS AND IMPLICATIONS FOR DECISION OR POLICY MAKING

The studies identified for this review investigated screening tools for the assessment of NP. Accuracy was variable depending on the clinical population assessed as well as the tool applied. The DN4 appears to perform well in certain populations (i.e., cancer, diabetes, breast resection surgery, leprosy, general chronic pain, and spinal cord injury), and to have higher accuracy than the LANSS, PD-Q, and Neuropathic Pain Questionnaire tools in some conditions (i.e., cancer, general chronic pain, spinal cord injury). Conversely, diagnostic accuracy outcomes in several patient populations (i.e., fibromyalgia, advanced metastatic cancer, failed back surgery syndrome, and neck and upper limb pain) should caution against the use of the screening tools identified in this review in these conditions. In addition to the questionnaires mentioned above, one four-step tool showed poor accuracy for general chronic pain, whereas a tool comprised of a cluster of model-derived signs and symptoms of NP showed promise in low back and leg pain.

Observed differences in the performance of various tools may be attributable to the absence of clinical assessment components within the PD-Q and Neuropathic Pain Questionnaire tools, as well as differences in the approach to clinical assessment taken for the LANSS and DN4. Also, the initial validation studies of the DN4[14] and LANSS[16] excluded patients with mixed pain types which could explain the poor performance of these tools in this context. These findings must also be interpreted in the context of the limitations of the individual studies, factoring in risk of bias, applicability, and generalizability, which was limited in some cases. Furthermore, the current best available option for clinical diagnosis is not without subjectivity due to dependence on clinical expertise, patient interviews, and inconsistent diagnostic and sensory tests. With the absence of an objective gold standard for the diagnosis of NP the diagnostic accuracy parameters covered in this report should be interpreted accordingly.

A European study reported that only 2% of chronic pain patients are treated by a pain specialist.[34] This implies that most patients with possible NP will be assessed at least initially by a primary care physician. Due to time constraints and the expertise required, the preferred IASP grading system may not be appropriate for use outside of secondary care. The screening and assessment tools captured within this report aim to empower primary care physicians without

comprehensive NP diagnostic training to participate in initial NP diagnosis, discern appropriateness of specialist referral, and make informed treatment decisions. The generally higher sensitivity (versus specificity) of the screening tools suggests that these tools may be appropriate for such uses. These screening tools are not intended to replace a comprehensive clinical diagnosis but rather to ensure the best possible diagnostic process when a pain specialist is not present in the care pathway. As such, the variable diagnostic accuracy of these screening tools within and between patient populations is not surprising.

In conclusion, the diagnostic accuracy of NP screening tools and assessment tools varies significantly depending on the tool and across patient populations. There is some evidence to suggest that the DN4 may be more universally applicable than other tools. When choosing a diagnostic approach, patient population, qualifications of the assessor, and clinical setting are important considerations.

PREPARED BY:

Canadian Agency for Drugs and Technologies in Health
Tel: 1-866-898-8439
www.cadth.ca

REFERENCES

1. Smith BH, Torrance N, Bennett MI, Lee AJ. Health and quality of life associated with chronic pain of predominantly neuropathic origin in the community. Clin J Pain. 2007 Feb;23(2):143-9.

2. Tarride JE, Collet JP, Choiniere M, Rousseau C, Gordon A. The economic burden of neuropathic pain in Canada. J Med Econ. 2006;9(1-4):55-68.

3. Treede RD, Jensen TS, Campbell JN, Cruccu G, Dostrovsky JO, Griffin JW, et al. Neuropathic pain: redefinition and a grading system for clinical and research purposes. Neurology. 2008 Apr 29;70(18):1630-5.

4. Bouhassira D, Letanoux M, Hartemann A. Chronic pain with neuropathic characteristics in diabetic patients: a French cross-sectional study. PLoS ONE [Internet]. 2013 Sep 13 [cited 2015 Apr 7];8(9):e74195. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3772849

5. Baron R. Mechanisms of disease: neuropathic pain--a clinical perspective. Nat Clin Pract Neurol [Internet]. 2006 Feb [cited 2015 Mar 24];2(2):95-106. Available from: http://www.nature.com/nrneurol/journal/v2/n2/full/ncpneuro0113.html

6. Rayment C, Hjermstad MJ, Aass N, Kaasa S, Caraceni A, Strasser F, et al. Neuropathic cancer pain: prevalence, severity, analgesics and impact from the European Palliative Care Research Collaborative-Computerised Symptom Assessment study. Palliat Med. 2013 Sep;27(8):714-21.

7. Jespersen A, Amris K, Bliddal H, Andersen S, Lavik B, Janssen H, et al. Is neuropathic pain underdiagnosed in musculoskeletal pain conditions? The Danish PainDETECTive study. Curr Med Res Opin. 2010 Aug;26(8):2041-5.

8. Torrance N, Smith BH, Bennett MI, Lee AJ. The epidemiology of chronic pain of predominantly neuropathic origin. Results from a general population survey. J Pain. 2006 Apr;7(4):281-9.

9. Bouhassira D, Lanteri-Minet M, Attal N, Laurent B, Touboul C. Prevalence of chronic pain with neuropathic characteristics in the general population. Pain. 2008 Jun;136(3):380-7.

10. Toth C, Lander J, Wiebe S. The prevalence and impact of chronic pain with neuropathic pain symptoms in the general population. Pain Med. 2009 Jul;10(5):918-29.

11. Kehlet H, Jensen TS, Woolf CJ. Persistent postsurgical pain: risk factors and prevention. Lancet. 2006 May 13;367(9522):1618-25.

12. Haroutiunian S, Nikolajsen L, Finnerup NB, Jensen TS. The neuropathic component in persistent postsurgical pain: a systematic literature review. Pain. 2013 Jan;154(1):95-102.

13. Cruccu G, Truini A. Tools for assessing neuropathic pain. PLoS Med [Internet]. 2009 Apr 7 [cited 2015 Mar 24];6(4):e1000045. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2661248

14. Krause SJ, Backonja MM. Development of a neuropathic pain questionnaire. Clin J Pain. 2003 Sep;19(5):306-14.

15. Freynhagen R, Baron R, Gockel U, Tolle TR. painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Curr Med Res Opin. 2006 Oct;22(10):1911-20.

16. Bennett M. The LANSS Pain Scale: the Leeds assessment of neuropathic symptoms and signs. Pain. 2001 May;92(1-2):147-57.

17. Scholz J, Mannion RJ, Hord DE, Griffin RS, Rawal B, Zheng H, et al. A novel tool for the assessment of pain: validation in low back pain. PLoS Med [Internet]. 2009 Apr 7 [cited 2015 Mar 24];6(4):e1000047. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2661253

18. Van Seventer R, Vos C, Meerding W, Mear I, Le Gal M, Bouhassira D, et al. Linguistic validation of the DN4 for use in international studies. Eur J Pain. 2010 Jan;14(1):58-63.

19. Attal N, Lanteri-Minet M, Laurent B, Fermanian J, Bouhassira D. The specific disease burden of neuropathic pain: results of a French nationwide survey. Pain. 2011 Dec;152(12):2836-43.

20. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011 Oct 18;155(8):529-36.

21. Abdallah FW, Morgan PJ, Cil T, Escallon JM, Semple JL, Chan VW. Comparing the DN4 tool with the IASP grading system for chronic neuropathic pain screening after breast tumor resection with and without paravertebral blocks: a prospective 6-month validation study. Pain. 2015 Apr;156(4):740-9.

22. Markman JD, Kress BT, Frazer M, Hanson R, Kogan V, Huang JH. Screening for neuropathic characteristics in failed back surgery syndromes: challenges for guiding treatment. Pain Med. 2015 Mar;16(3):520-30.

23. Bryce TN, Richards JS, Bombardier CH, Dijkers MP, Fann JR, Brooks L, et al. Screening for neuropathic pain after spinal cord injury with the spinal cord injury pain instrument (SCIPI): a preliminary validation study. Spinal Cord. 2014 May;52(5):407-12.

24. Mick G, Baron R, Correa-Illanes G, Hans G, Mayoral V, Frias X, et al. Is an easy and reliable diagnosis of localized neuropathic pain (LNP) possible in general practice? Development of a screening tool based on IASP criteria. Curr Med Res Opin. 2014 Jul;30(7):1357-66.

25. Haroun OM, Hietaharju A, Bizuneh E, Tesfaye F, Brandsma JW, Haanpaa M, et al. Investigation of neuropathic pain in treated leprosy patients in Ethiopia: a cross-sectional study. Pain. 2012 Aug;153(8):1620-4.

26. Smart KM, Blake C, Staines A, Thacker M, Doody C. Mechanisms-based classifications of musculoskeletal pain: part 2 of 3: symptoms and signs of peripheral neuropathic pain in patients with low back (+/- leg) pain. Man Ther. 2012 Aug;17(4):345-51.

27. Sadler A, Wilson J, Colvin L. Acute and chronic neuropathic pain in the hospital setting: use of screening tools. Clin J Pain. 2013 Jun;29(6):507-11.

28. Gauffin J, Hankama T, Kautiainen H, Hannonen P, Haanpaa M. Neuropathic pain and use of PainDETECT in patients with fibromyalgia: a cohort study. BMC Neurol [Internet]. 2013 Feb 14 [cited 2015 Mar 9];13:21. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3582578

29. Spallone V, Morganti R, D'Amato C, Greco C, Cacciotti L, Marfia GA. Validation of DN4 as a screening tool for neuropathic pain in painful diabetic polyneuropathy. Diabet Med. 2012 May;29(5):578-85.

30. Lasry-Levy E, Hietaharju A, Pai V, Ganapati R, Rice AS, Haanpää M, et al. Neuropathic pain and psychological morbidity in patients with treated leprosy: a cross-sectional prevalence study in Mumbai. PLoS Negl Trop Dis [Internet]. 2011 Mar 8 [cited 2015 Mar 10];5(3):e981. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3050898

31. Pérez C, Sánchez-Martinez N, Ballesteros A, Blanco T, Collazo A, González F, et al. Prevalence of pain and relative diagnostic performance of screening tools for neuropathic pain in cancer patients: a cross-sectional study. Eur J Pain. 2014 Sep 30.

32. Tampin B, Briffa NK, Goucke R, Slater H. Identification of neuropathic pain in patients with neck/upper limb pain: application of a grading system and screening tools. Pain. 2013 Dec;154(12):2813-22.

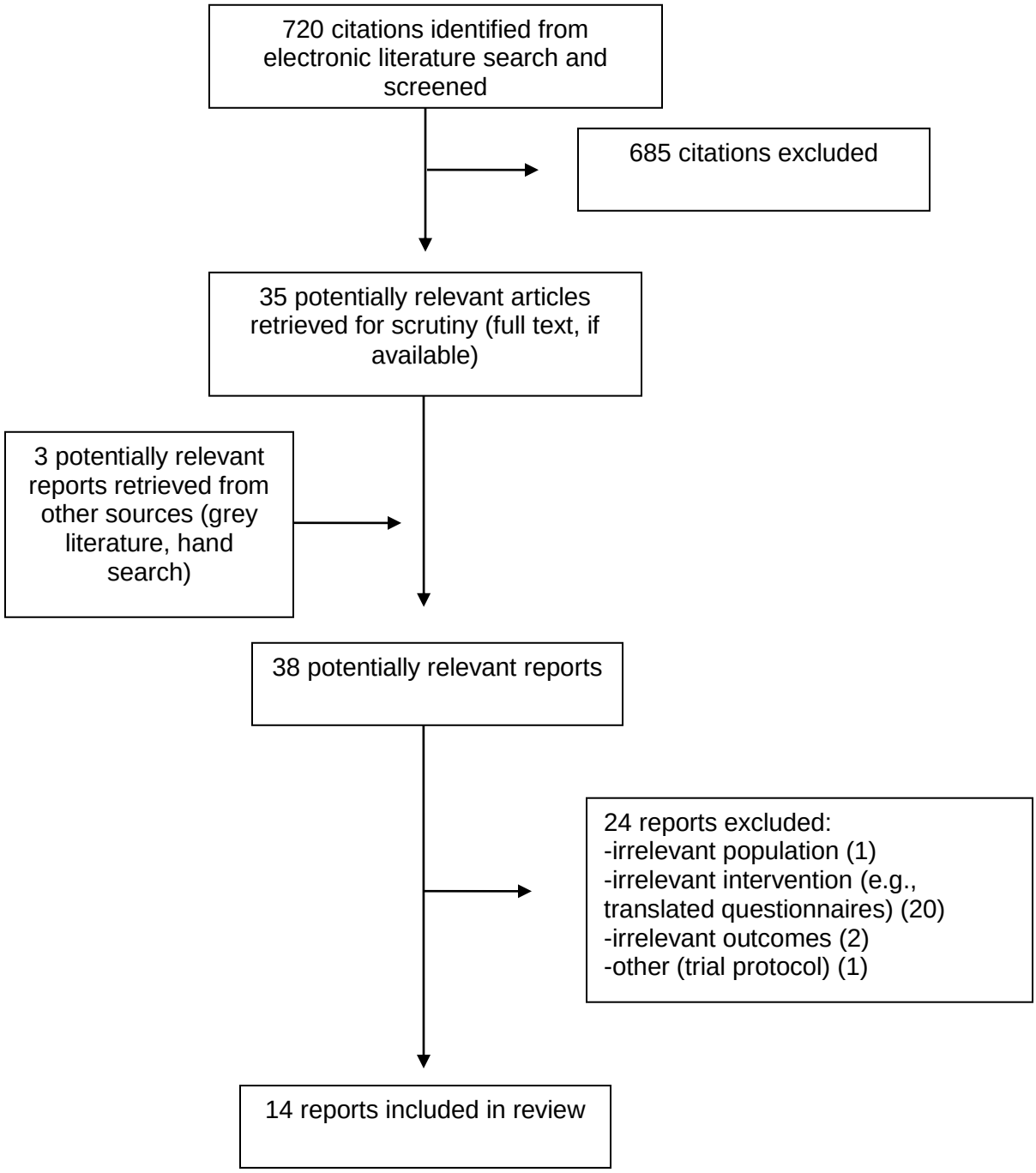
33. Hallstrom H, Norrbrink C. Screening tools for neuropathic pain: can they be of use in individuals with spinal cord injury? Pain. 2011 Apr;152(4):772-9.

34. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain. 2006 May;10(4):287-333.

35. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol. 2009 Oct;62(10):e1-e34.

36. Cho SI, Lee CH, Park GH, Park CW, Kim HO. Use of S-LANSS, a tool for screening neuropathic pain, for predicting postherpetic neuralgia in patients after acute herpes zoster events: a single-center, 12-month, prospective cohort study. J Pain. 2014 Feb;15(2):149-56.

APPENDIX 1: Selection of Included Studies



APPENDIX 2: Characteristics of Included Publications

Table A1: Characteristics of Included Diagnostic Accuracy Studies

First Author, Publication Year, Country	Study Design	Patient Characteristics, Sample Size	Setting	Index Test(s)	Reference Test(s)	Outcomes
Abdallah, 2015, Canada[35]	Cross sectional (included in prospective follow-up to RCT)	Females post-breast tumor resection with and without paravertebral blocks, n = 66	Primary care hospital	DN4 *(performed by research assistant)*	IASP Grading System *(performed by research assistant)*	Sensitivity, specificity, AUROC
Markman, 2014, US[36]	Cross sectional study (included in retrospective chart review)	Patients with failed back surgery syndrome, n = 44	Secondary care (research center)	DN4, LANSS	Quebec Task Force Classification of Spinal Disorders	Sensitivity, specificity
Mick, 2014, France[24]	Cross sectional (within prospective cohort study	Patients with chronic pain, n = 360	Primary care	4 question screening tool and algorithm including sensory examination based on IASP criteria conducted by 31 general practitioners DN4 conducted by three pain specialists	Clinical diagnosis by three pain specialists using IASP criteria	Sensitivity, specificity, NPV, PPV
Bryce, 2014, US[23]	Cross-sectional (embedded in larger multicenter RCT)	Patients with spinal cord injury being treated for depression, n = 82 pain sites (n = 36 subjects)	Spinal cord injury centres	Spinal Cord Injury Pain Instrument (SCIPI) *(administered by research assistant)*	Clinical determination of pain using IASP grading scheme (certainty rating 4 and 5 only)	Sensitivity, specificity, AUROC
Perez, 2014, Spain[31]	Cross-sectional	Patients undergoing chemotherapy for any cancer, n = 359	Oncology unit	LANSS, DN4, PD-Q	Clinical diagnosis by pain specialist based on IASP grading system *(patients with neurotoxic exposure*	Sensitivity, specificity

Table A1: Characteristics of Included Diagnostic Accuracy Studies						
First Author, Publication Year, Country	**Study Design**	**Patient Characteristics, Sample Size**	**Setting**	**Index Test(s)**	**Reference Test(s)**	**Outcomes**
					were graded probable for NP, and patients with further diagnostic confirmation of lesion were graded definite)	
Sadler, 2013, UK[27]	Cross-sectional	Outpatients attending a pain service clinic, n = 67	Chronic pain center or hospital	Self-reported LANSS and DN4	Classification as NP, mixed, or non-NP by pain specialist	Sensitivity, specificity, PPV, NPV
Tampin, 2013, Australia	Cross-sectional	Primary care patients with neck and upper limb pain with a suspected nerve lesion, n = 152	Hospital neurosurgery triage clinic	PD-Q (patient administered) and LANSS (administered unblinded at the end of clinical examination) questionnaires	Clinical grading system (IASP – NeuPSIG) based on history, examination results (neurological and musculoskeletal status) and results of investigations	Sensitivity, specificity, PPV, NPV, LR, DOR, AUROC
Gauffin, 2013, Finland[28]	Cross-sectional	Patients with fibromyalgia, n = 156	Outpatient clinic (secondary health care setting)	PD-Q	Clinical assessment by single physician (clinical history, physician examination, neurological examination, sensory testing) by European Federation of Neurological Societies guidelines (2004)	Sensitivity, specificity, PPV, LR, Youden's index, AUROC
Rayment, 2012, International[6]	Cross-sectional (multicenter)	Patients with incurable metastatic or locally advanced cancer, n = 560	Palliative care and hospital oncology wards	PD-Q	Clinical diagnosis recorded on the Edmonton Classification System for Cancer Pain (ECS-SP)	Sensitivity, specificity, PPV, NPV
Haroun, 2012, Ethiopia[25]	Cross-sectional	Patients with leprosy who had completed multi-drug therapy, n = 80	Hospitals and health centres	DN4 and LANSS	Clinical assessment based on whether distribution of pain was not neuroanatomically possible, and whether clinical	Sensitivity, specificity

Table A1: Characteristics of Included Diagnostic Accuracy Studies						
First Author, Publication Year, Country	**Study Design**	**Patient Characteristics, Sample Size**	**Setting**	**Index Test(s)**	**Reference Test(s)**	**Outcomes**
					examination confirmed negative or positive sensory signs were confined to innervation territory of lesioned nervous structure	
Smart, 2012, Ireland[26]	Cross-sectional	Patients with low back and leg pain presenting from physiotherapy, n = 452	Four Hospitals	Cluster of two symptoms and one sign predictive of PNP *(i.e., pain referred in a dermatomal or cutaneous distribution, history of nerve injury pathology or mechanical compromise, pain/symptom provocation with mechanical/movement tests (e.g. active/passive, neurodynamic) that move/load/compress neural tissue)*	Mechanisms-based classification based on experienced clinical judgement	Sensitivity, specificity, DOR
Spallone, 2011, Italy[29]	Cross-sectional	Patients with diabetes, n = 158	Outpatient clinic	DN4 and DN4 interview (second investigator blinded to results)	Clinical Assessment (by single investigator) Scoring symptom for symptoms and signs, quantitative sensory testing, nerve conduction studies, pain history, numerical rating scale, and Short-Form McGill Pain Questionnaire	Sensitivity, specificity, PPV, NPV, LR, AUROC
Lasry-Levy, 2011, India[30]	Cross-sectional	Patients with leprosy, n = 101	Outpatient clinics (resource poor field	DN4 (translated into Hindi/Marathi)	Clinical diagnosis[3] by clinician (distribution of pain anatomically plausible, confirmation tests of	Sensitivity, specificity

Table A1: Characteristics of Included Diagnostic Accuracy Studies

First Author, Publication Year, Country	Study Design	Patient Characteristics, Sample Size	Setting	Index Test(s)	Reference Test(s)	Outcomes
			setting)		neurological examination demonstrated positive or negative sensory signs confined to innervation territory of affected peripheral nerves, pain associated with having leprosy)	
Hallstrom, 2011, Sweden[33]	Cross-sectional	Patients with spinal cord injury, n = 40	Spinal cord unit within Hospital	DN4, LANSS, NPQ, PD-Q (in Swedish) All with the addition of 5 yes/no questions that could be of importance for discriminating NP from non-NP (based on clinical opinion)	Clinical diagnosis based on European Federation of Neurological Societies guidelines	Sensitivity, specificity

AUROC = area under the receiver operating characteristic curve; DN4 = Douleur Neuropathique en 4 questions; DOR = diagnostic odds ratio; IASP = International Association for the Study of Pain; LANSS = Leeds Assessment of Neuropathic Symptoms and Signs; LR = likelihood ratio; NeuPSIG = Neuropathic Pain Special Interest Group; NP = neuropathic pain; PD-Q = painDETECT questionnaire; PPV = positive predictive value; RCT = randomized controlled trial; UK = United Kingdom; US = United States

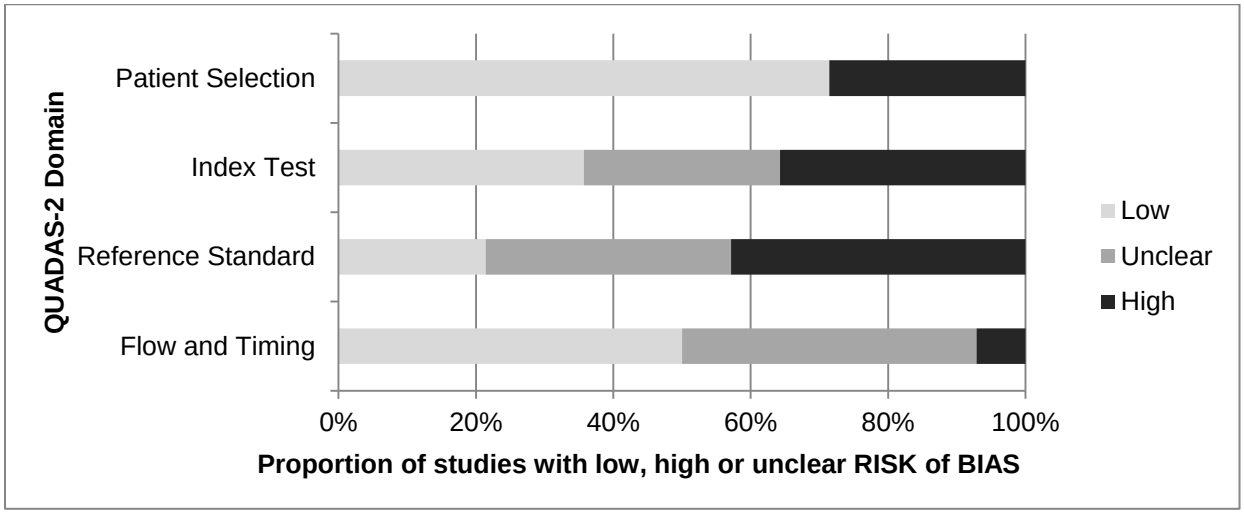
APPENDIX 3: Critical Appraisal of Included Publications

Table A2. QUADAS-2 Ratings of Included Studies

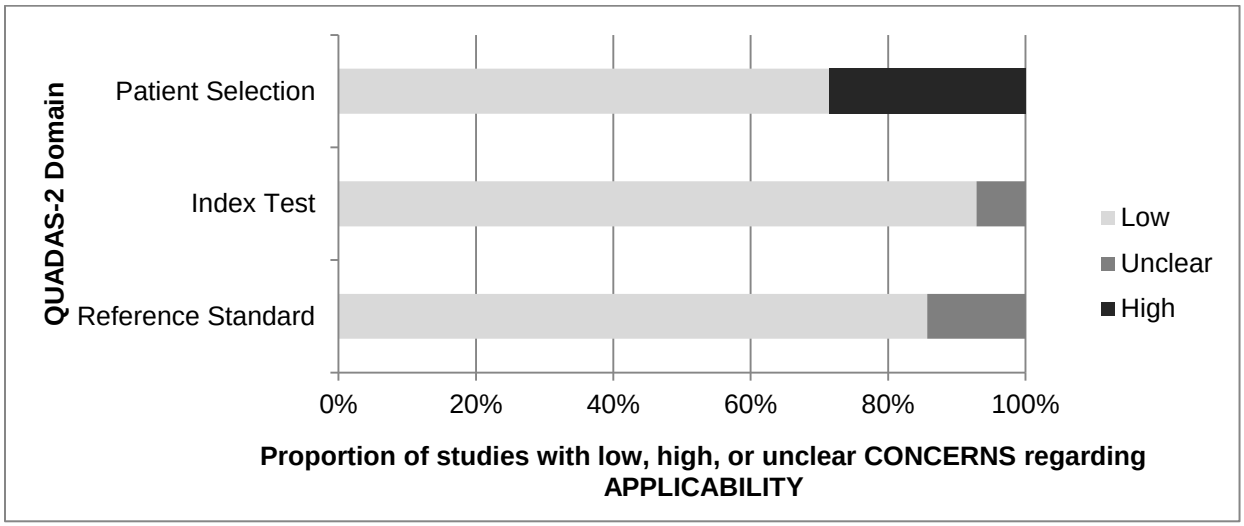
Study	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Abdallah[35]	⊕	⊖	⊖	⊕	⊕	⊕	⊕
Markman[36]	⊕	⊕	⊕	?	⊕	?	⊕
Mick[24]	⊕	⊕	⊕	⊖	⊕	⊕	⊕
Bryce[23]	⊖	⊖	⊖	?	⊖	⊕	⊕
Perez[31]	⊖	⊕	⊖	⊕	⊕	⊕	⊕
Sadler[27]	⊕	⊖	⊖	?	⊕	⊕	⊕
Tampin[32]	⊕	⊖	⊕	⊕	⊕	⊕	⊕
Gauffin[28]	⊕	?	?	?	⊖	⊕	⊕
Rayment[6]	⊖	?	?	⊕	⊖	⊕	⊕
Haroun[25]	⊕	?	?	?	⊕	⊕	⊕
Smart	⊖	⊖	?	?	⊖	⊕	?
Spallone[29]	⊕	⊕	⊖	⊕	⊕	⊕	⊕
Lasry-Levy[30]	⊕	?	?	⊕	⊕	⊕	?
Hallstrom[33]	⊕	⊕	⊖	⊕	⊕	⊕	⊕

Legend: ⊕ low risk; ⊖ high risk; ? = unclear

Figure A1. Graphical Display of QUADAS-2 Results by Proportion of Low, High, or Unclear Bias or Applicability



1. Risk of Bias



2. Applicability

Table A3: Specific Strengths and Limitations of Included Studies using QUADAS-2

Strengths	Limitations
Abdallah[35]	
Risk of Bias • Consecutive sample of participants enrolled • Case-control design avoided • Threshold of tests was pre-specified • Reference test is regarded as the clinical standard • Reference assessor was blinded • Index and reference tests were conducted during the same period	Risk of Bias • Included individuals who had undergone paravertebral blocks and those who had not • Single assessor completed both the reference and index tests • Index test may have been completed with knowledge of reference test results • Reference test was missing diagnostic component, limiting utility • Assessment tests were conducted by a research

Table A3: Specific Strengths and Limitations of Included Studies using QUADAS-2	
Strengths	**Limitations**
• All patients received the same reference standard • Minimal losses to follow up Applicability • Patient population appropriate for research question • Index and reference tests specific for measuring NP	assistant trained in application of the tools but with unclear clinical credentials (i.e., not conducted by pain specialist) • Blinding of index assessor unclear • One practitioner conducted all the assessments Applicability • Patients only representative of individuals undergoing breast tumor resection
Markman[36]	
Risk of Bias • Consecutive sample of patients enrolled • Case-control design avoided • No inappropriate patient exclusion • All patients received the same reference standard • Blinded investigator conducted reference tests • All patients included in the analysis • Minimal loss to follow up Applicability • Patients and index test appropriate for research question	Risk of Bias • Patient flow unclear • Patient information was retrieved retrospectively, possible recall bias • Pre-specification of test thresholds was unclear • Diagnostic test administrator unclear for both index and reference tests • Patients only representative of individuals with failed back surgery syndrome Applicability • Appropriateness of reference standard for NP unclear
Mick[24]	
Risk of Bias • Consecutive sample of patients enrolled • Case-control design avoided • No inappropriate patient exclusion • Good generalizability of tool accuracy given chronic pain population was unselected – no exclusions for pain of assumed mixed origin • Index test performed prior to reference standard • Threshold for index test pre-specified (algorithm of 4 questions) • Reference standard appropriate for target condition • Reference test assessor blinded • Same reference standard assessed in all patients referred for assessment by pain specialist • Statistical corrections were made for incomplete verification and dropouts Applicability • Patients and setting match those established in the review question • Conduct of index and reference tests appropriate for the review question	Risk of Bias • Up to 14 day lag between index and reference test • Reference test not completed on all individuals who underwent index test (selected for more probable NP) • Tests completed by multiple practitioners (31 for index, 3 for reference) • Tests completed in different settings (primary care versus hospital) • Only every 10th patient without LNP was referred for reference test – all patients with LNP referred • Method of recruitment unclear • Diagnostic accuracy was a secondary outcome • A proportion of referred patients did not visit the specialist and 1/10 patients with diagnosis of nLNP or nNP were allocated to specialist • Not all patients were included in the analysis – filtered out to select for more patients with probable NP diagnosis – reduced generalizability to wider chronic pain population, especially those with possible mixed pain origin
Bryce[23]	
Risk of Bias • Consecutive patients enrolled in multicenter	Risk of Bias • Difficult to diagnose patients were excluded

Table A3: Specific Strengths and Limitations of Included Studies using QUADAS-2	
Strengths	**Limitations**
RCT for treatment of depression with venlafaxine hydrochloride • Case-control design avoided • Only individuals undergoing alternate treatment for depression, with suicidal intent or plan, substance dependence, or on non-stable doses of psychoactive medications were excluded • Clinical assessors were blinded to index test results • Reference test appropriate for target condition Applicability: • Index test representative of intended assessment	(only patients with 4-5 certainty rating in clinical assessment included) • All patients had at least moderate depression – limited generalizability • Clinical classifications were only of high confidence – those with low confidence excluded – diagnostic accuracy of SCIPI may be lower in all assessed patients • Assessments conducted by multiple clinicians and research assistants • Threshold was not pre-specified – thresholds to optimize sensitivity and specificity were proposed post-hoc • Only highly certain reference standard assessments were considered plausible • Patient flow unclear • Multiple pain sites assessed independently for single patients Applicability: • Patients may not be representative of all patients with spinal cord injury as only those with certain clinical diagnoses of NP were included in analysis
Perez[31]	
Risk of Bias • Case-control design avoided • No inappropriate patient exclusions • Index test assessors blind to reference test results • Thresholds for index tests were pre-specified • Index and reference test assessments occurred during the same visit Applicability • Included patients, index and tests match review question	Risk of Bias • Recruitment was ad hoc based on availability of investigator • Single-center study • Only some patients had diagnostic imaging or neurophysiologic tests in their medical files Applicability • Reference test did not include comprehensive diagnostic tests (excluded quantitative sensory testing) resulting in lacking confidence in classification of individuals
Sadler[27]	
Risk of Bias • Consecutive sample of patients enrolled • Case-control design avoided • Thresholds were established a priori • Reference assessor was blinded for pain service patients • Index and reference tests were completed on the same day • All patients were included in analysis Applicability • Included patients, reference, and index test match review question	Risk of Bias • Method of classification for reference did not follow most recent guidelines despite being available at time of publication • Both outpatients and surgical populations were included – reason for inclusion was unclear • Index test was conducted by different pain specialists depending on the patient group and site • Patient flow unclear • Multiple pain specialists classified NP (reference test)

Table A3: Specific Strengths and Limitations of Included Studies using QUADAS-2	
Strengths	**Limitations**
Tampin[32]	
Risk of Bias • Consecutive sample of patients enrolled • Case-control design avoided • Reference standard appropriate for target condition • Reference standard conducted without knowledge of index test results • All assessments were completed on the same day Applicability • Target condition and reference standard match research question	Risk of Bias • One index test (PD-Q) performed before clinical evaluation and one performed after (LANSS) • Clinical assessor was blinded to PD-Q results • LANSS assessor was not blinded • LANSS was assessed by the same practitioner who completed the clinical assessment (reference standard) • Cut-off scores were based on ROC curve analysis – cut-offs may only be appropriate for this study population • Index tests were not conducted under the same circumstances (one before clinical exam, one after) • Patients had been referred by general practitioner • LANSS index test conducted with knowledge of clinical findings
Gauffin[28]	
Risk of Bias • A consecutive sample of patients was enrolled (targeted retrospective sample of FM patients) • Case control design avoided • Index test appropriate for review question Applicability • Due to exclusion of patients previously diagnosed with NP patient population may have less prevalent and severe NP • Reference standard and index tests were appropriate for research question	Risk of Bias • Blinding of test assessors was unclear • Threshold was optimized for sensitivity and specificity (although pre-specified thresholds were assessed) • Self-administration of index test may introduce subjectivity • Reference standard (clinical assessment) did not follow the most up to date guidelines • Duration between index and reference tests were unclear Applicability • Patient population excluded patients who had been previously diagnosed with NP
Rayment[6]	
Risk of Bias • Consecutive sample of patients enrolled • Case-control design avoided • Index and reference test data collected during the same time period Applicability • Index and tests were appropriate for review question	Risk of Bias • Standardization of reference standard across sites was unclear • Unclear differences in patient characteristics for sub-population for which PD-Q data was available • Blinding of reference and index assessors was unclear • Reference test was not up to date with most recent clinical guidelines Applicability • Population may not represent those patients with more advanced disease who would be unfit to complete assessments (excluded those with more advanced disease)

Table A3: Specific Strengths and Limitations of Included Studies using QUADAS-2	
Strengths	**Limitations**
Haroun[25]	
Risk of Bias • Consecutive sample of patients recruited • Case-control design avoided • Test thresholds were pre-specified • Same reference standard applied to all patients • No loss to follow up Applicability • Patients and setting appropriate for review question • Index test appropriate for review question	Risk of Bias • Patients with comorbidities that could contribute to NP (e.g., diabetes) excluded (severely ill patients excluded) • Blinding of investigators to index and reference results unclear • Interval between index and reference tests unclear Applicability • Reference test did not follow most recent clinical guidelines for assessing NP
Smart[26]	
Risk of Bias • Case-control design avoided • Criteria for NP • All patients received the same reference standard Applicability • Patients appropriate for review question	Risk of Bias • Convenience sample • Index method determined using logistic regression to identify clusters of signs and symptoms • Criteria for index test developed through modeling methods • Patients with mixed or indiscriminate pain were excluded (hard to diagnose patients) • The same clinician completed the reference and index test assessments • Time of administration of screening unclear • Reference method did not follow most recent guidelines for NP diagnosis Applicability • Appropriateness of index and reference methods for NP unclear
Spallone[29]	
Risk of Bias • Consecutive sample of patients enrolled • Case-control design avoided • Single investigators responsible for both reference and index tests • Both index and reference test investigators blinded • Thresholds for index test were pre-specified • Both assessments completed on the same day • All patients received the same reference standard Applicability • Included patients are appropriate for review question • Both index and reference tests were appropriate for review question	Risk of Bias • Patients with NP due to non-diabetes related causes were excluded (only representative of patients with pain related to diabetic neuropathy) • Reference assessment did not follow most recent guidelines for identification of NP

Table A3: Specific Strengths and Limitations of Included Studies using QUADAS-2	
Strengths	**Limitations**
Lasry-Levy[30]	
Risk of Bias • Consecutive sample of patients enrolled • Case-control design avoided • No loss to follow up Applicability • Included patients are appropriate for review question • Index test and target condition as defined by the reference test appropriate	Risk of Bias • Patients who had not undergone multi-drug therapy were excluded • Patients who had other potential pathologies related to NP were not excluded • Blinding of assessors was unclear • Reference assessment did not follow most recent guidelines for identification of NP (may not have been available at time of assessment) • Interval between index and reference tests unclear • Credentials of reference test administrator unclear • Process of reference test unclear • Threshold specification unclear
Hallstrom[33]	
Risk of Bias • Consecutive sample of patients enrolled • Case-control design avoided • Cut-off thresholds were pre-specified for reference and index tests • Interviewers and clinical assessors were blinded for index and reference assessments • Questionnaires were administered in a random order • All patients received same reference standard • Minimal losses to follow up • All patient assessments were done on a single day Applicability • Included patients, index test, reference test were appropriate for review question	Risk of Bias • Two physicians administered the reference test • Limited sample size • Reference assessment did not follow most recent guidelines for identification of NP (may not have been available)

FM = fibromyalgia; LANSS = Leeds Assessment of Neuropathic Symptoms and Signs; LNP = localized neuropathic pain; nLNP = non-localized neuropathic pain; NP = neuropathic pain; PD-Q = painDETECT questionnaire; RCT = randomized controlled trial; SCIPI = spinal cord injury pain assessment

APPENDIX 4: Main Study Findings and Author's Conclusions

Table A4: Summary of Findings of Included Studies

Author	Accuracy Measures (95% CIs)	Author's Conclusions
Abdallah[35]	DN4: Sensitivity = 90% Specificity = 60% AUROC = 0.771 (0.627 to 0.914)	• The DN4 has good research and clinical utility for identification of CNP following breast tumor resection • Low specificity may be attributable to either limitations in the IASP criteria for ruling out CNP, rather than intrinsic DN4 limitations
Markman[36]	DN4: Sensitivity = 62% Specificity = 44% LANSS: Sensitivity = 38% Specificity = 75% Combined Questionnaires: Sensitivity = 20% Specificity = 58% PPV = 16.7%	• Neuropathic screening questionnaires have low diagnostic accuracy for identifying signs and symptoms of NP in failed back surgery syndrome patients • Variability in rates of NP between tools may be attributed to different weighting of each tool's sensory and effective terminology
Mick[24]	4 question screening tool: Sensitivity = 51.6% Specificity = 79.0% PPV = 51.4% NPV = 79.2% DN4: Sensitivity = 87.1% Specificity = 57.4% PPV = 47.9% NPV = 90.9%	• The 4 question screening tool performs well for exclusion of LNP as a diagnosis and is appropriate for diagnosis of NP in chronic pain patients in primary care • May be a fast alternative to other more complex assessment methods • Screening tool had greater specificity, whereas the DN4 exhibited higher sensitivity relative to clinical diagnosis using IASP grading criteria in the diagnosis of NP • Both tools had comparable predictive value • DN4 was only conducted in a selected pain population (who were already screened which may overestimate diagnostic accuracy/generalizability
Bryce[23]	SCIPI: Sensitivity = 72% Specificity = 78% AUROC = 0.77	• SCIPI can be used to reliably distinguish neuropathic from non-NP • To maximize sensitivity a cut-off of 1 may be appropriate, if specificity is more of a concern a cut-off of 2 may be appropriate
Perez[31]	DN4 Sensitivity = 88% Specificity = 88% LANSS Sensitivity = 68% Specificity = 93% PD-Q Sensitivity = 18% Specificity = 98%	• DN4 was the most balanced in terms of sensitivity and specificity • PD-Q had greater specificity than DN4 at the expense of sensitivity • LANSS had greater specificity than DN4 at expense of sensitivity • PD-Q had very low sensitivity, despite good specificity. May be due to its initial development for patients with low back pain
Sadler[27]	LANSS Sensitivity = 76% Specificity = 94%	• Sensitivity of LANSS and DN4 was similar • LANSS had higher specificity for diagnosing NP in this chronic pain and surgical population.

Table A4: Summary of Findings of Included Studies		
Author	**Accuracy Measures (95% CIs)**	**Author's Conclusions**
	PPV = 0.87 NPV = 0.89 <u>DN4</u> Sensitivity = 76% Specificity = 70% PPV = 0.57 NPV = 0.85	
Tampin[32]	<u>LANSS</u> Sensitivity = 22% Specificity = 88% PPV = 0.44 NPV = 0.31 LR+ = 1.83 LR- = 0.89 DOR = 2.0 AUROC = 0.73 (0.64 to 0.81) <u>PD-Q</u> Sensitivity = 64% Specificity = 62% PPV = 0.42 NPV = 0.80 LR+ = 1.68 LR- = 0.58 DOR = 2.9 AUROC = 0.63 (0.53 to 0.73)	• Both questionnaires had limited ability to identify patients with neck and upper limb pain with clinically defined NP • Higher sensitivity demonstrated by PD-Q and higher specificity demonstrated by LANSS • Sensitivity of LANSS was substantially lower than other accuracy measures
Gauffin[28]	<u>PD-Q:</u> Sensitivity = 79% Specificity = 53% PPV = 0.46 (0.36 to 0.57) LR+ = 1.7 (1.33 to 2.17) AUROC = 0.69 (0.60 to 0.77) Youden's index = 17	• PD-Q does not demonstrate accurate diagnostic capacity for identifying NP in patients with fibromyalgia
Rayment[6]	<u>PD-Q:</u> Sensitivity = 53% Specificity = 77% PPV = 0.33 NPV = 0.89	• Sensitivity of PD-Q is poor for patients with incurable or locally advanced cancer • NP screening tools may need to be developed and adapted for this clinical population
Haroun[25]	<u>DN4:</u> Sensitivity = 100% Specificity = 45% <u>LANSS:</u> Sensitivity = 85% Specificity = 42%	• DN4 had higher sensitivity than the LANSS though both tools performed well in the leprosy population • Specificity of both tools was significantly lower suggesting that there may be tendency for patients to rate non NP pain as NP

Table A4: Summary of Findings of Included Studies		
Author	**Accuracy Measures (95% CIs)**	**Author's Conclusions**
Smart[26]	Model Specified Cluster of Symptoms and Signs: Sensitivity = 0.86 (0.78 to 0.92) Specificity = 0.96 (0.93 to 0.98) PPV = 0.86 (0.78 to 0.92) LR+ = 21.6 (12.8 to 36.2) DOR = 150.9 (69.4 to 328.1)	• The model-derived cluster of symptoms and signs had good sensitivity and specificity for identifying NP as defined by expert clinical judgement • This diagnostic approach could be useful in the clinical population of patients seeking physiotherapy with reasonable physical function in whom assumed dominance of peripheral NP is assumed
Spallone[29]	DN4: Sensitivity = 80% Specificity = 92% PPV = 0.82 NPV = 0.91 LR+ = 9.6 LR- = 0.22 AUROC = 0.94 (0.90 to 0.97) DN4 Interview: Sensitivity = 84% Specificity = 84% PPV = 0.71 NPV = 0.92 LR+ = 5.3 LR- = 0.19 AUROC = 0.93 (0.89 to 0.96)	• Both the full DN4 and its questionnaire only (no bedside physical examination) performed well compared to clinical examination for accurately identifying NP in patients with diabetic neuropathy • DN4 Interview had slightly lower specificity than the DN4 Full suggesting it could be used well as a first assessment in clinical practice • DN4 fails to identify ~20% of patients with NP – therefore, doesn't exclude presence of condition but offers guidance for further diagnostic evaluation
Lasry-Levy[30]	DN4: Sensitivity = 100% Specificity = 92%	• DN4 has perfect sensitivity and high specificity for identifying NP cases in leprosy patients assessed in a resource poor field setting
Hallstrom[33]	DN4: Sensitivity = 93% Specificity = 75% AUROC = 0.86 LANSS: Sensitivity = 36% Specificity = 100% AUROC = 0.81 NPQ: Sensitivity = 50% Specificity = 100% AUROC = 0.79 PD-Q: Sensitivity = 68%	• The screening tools assessed showed varied goodness of fit with the DN4 having the best overall results based on sensitivity, specificity, and AUROC • Only the DN4 is reliable for identifying NP in the population of patients with SCI and pain • Specificity of the LANSS and NPQ failed to discriminate more than 50% of those with NP, decreasing their suitability for assessing patients with SCI

Table A4: Summary of Findings of Included Studies		
Author	**Accuracy Measures (95% CIs)**	**Author's Conclusions**
	Specificity = 83% AUROC = 0.71	

AUROC = area under the receiver operating characteristic curve; CNP = chronic neuropathic pain; DN4 = Douleur Neuropathique en 4 questions; DOR = diagnostic odds ratio; IASP = International Association for the Study of Pain; LANSS = Leeds Assessment of Neuropathic Symptoms and Signs; LR = likelihood ratio; NP = neuropathic pain; NPQ = neuropathic pain questionnaire; NPV = negative predictive value; PD-Q = painDETECT questionnaire; PPV = positive predictive value; SCIPI = Spinal Cord Injury Pain Instrument

APPENDIX 5: Additional References of Potential Interest

Validation and Diagnostic Accuracy of Translated Questionnaires

1. Turkel Y, Turker H, Demir IA, Bayrak AO, Onar MK. Validation of self report version of the Leeds Assessment of Neuropathic Symptoms and Signs score for identification of neuropathic pain in patients from northern Turkey. Adv Clin Exp Med [Internet]. 2014 Jul-Aug [cited 2015 Mar 19];23(4):599-603. Available from: http://www.advances.am.wroc.pl/pdf/2014/23/4/599.pdf

2. Sykioti P, Zis P, Vadalouca A, Siafaka I, Argyra E, Bouhassira D, et al. Validation of the Greek Version of the DN4 Diagnostic Questionnaire for Neuropathic Pain. Pain Pract. 2014 May 5.

3. Alkan H, Ardic F, Erdogan C, Sahin F, Sarsan A, Findikoglu G. Turkish version of the painDETECT questionnaire in the assessment of neuropathic pain: a validity and reliability study. Pain Med. 2013 Dec;14(12):1933-43.

4. Hamdan A, Luna JD, Del Pozo E, Galvez R. Diagnostic accuracy of two questionnaires for the detection of neuropathic pain in the Spanish population. Eur J Pain. 2014 Jan;18(1):101-9.

5. Madani SP, Fateh HR, Forogh B, Fereshtehnejad SM, Ahadi T, Ghaboussi P, et al. Validity and reliability of the persian (Farsi) version of the DN4 (Douleur Neuropathique 4 Questions) questionnaire for differential diagnosis of neuropathic from non-neuropathic pains. Pain Pract. 2014 Jun;14(5):427-36.

6. Erdemoglu AK, Koc R. Brief Pain Inventory score identifying and discriminating neuropathic and nociceptive pain. Acta Neurol Scand. 2013 Nov;128(5):351-8.

7. Adelmanesh F, Jalali A, Attarian H, Farahani B, Ketabchi SM, Arvantaj A, et al. Reliability, validity, and sensitivity measures of expanded and revised version of the short-form McGill Pain Questionnaire (SF-MPQ-2) in Iranian patients with neuropathic and non-neuropathic pain. Pain Med. 2012 Dec;13(12):1631-6.

8. Van Seventer R, Vos C, Giezeman M, Meerding WJ, Arnould B, Regnault A, et al. Validation of the Dutch version of the DN4 diagnostic questionnaire for neuropathic pain. Pain Pract. 2013 Jun;13(5):390-8.

9. Li J, Feng Y, Han J, Fan B, Wu D, Zhang D, et al. Linguistic adaptation, validation and comparison of 3 routinely used neuropathic pain questionnaires. Pain Physician. 2012 Mar-Apr;15(2):179-86.

10. De Andres J, Perez-Cajaraville J, Lopez-Alarcon MD, Lopez-Millan JM, Margarit C, Rodrigo-Royo MD, et al. Cultural adaptation and validation of the painDETECT scale into Spanish. Clin J Pain. 2012 Mar-Apr;28(3):243-53.

11. Sommer C, Richter H, Rogausch JP, Frettloh J, Lungenhausen M, Maier C. A modified score to identify and discriminate neuropathic pain: a study on the German version of the Neuropathic Pain Symptom Inventory (NPSI). BMC Neurol [Internet]. 2011 Aug 23 [cited

2015 Apr 7];11:104. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3180265

12. Chan A, Wong S, Chen PP, Tsoi TH, Lam J, Ip WY, et al. Validation study of the Chinese Identification Pain Questionnaire for neuropathic pain. Hong Kong Med J. 2011 Aug;17(4):297-300.

13. Kitisomprayoonkul W. Validation study of the Thai ID Pain Scale. J Med Assoc Thai. 2011 May;94(5):610-5.

14. Villoria J, Rodriguez M, Berro MJ, Stern A, Sanchez-Magro I. Psychometric validation of the neuropathic pain symptom inventory for its use in Spanish. J Pain Symptom Manage. 2011 Jul;42(1):134-46.

15. Harifi G, Ouilki I, El Bouchti I, Ouazar MA, Belkhou A, Younsi R, et al. Validity and reliability of the Arabic adapted version of the DN4 questionnaire (Douleur Neuropathique 4 Questions) for differential diagnosis of pain syndromes with a neuropathic or somatic component. Pain Pract. 2011 Mar-Apr;11(2):139-47.

16. Koc R, Erdemoglu AK. Validity and reliability of the Turkish Self-administered Leeds Assessment of Neuropathic Symptoms and Signs (S-LANSS) questionnaire. Pain Med. 2010 Jul;11(7):1107-14.

17. Unal-Cevik I, Sarioglu-Ay S, Evcik D. A comparison of the DN4 and LANSS questionnaires in the assessment of neuropathic pain: validity and reliability of the Turkish version of DN4. J Pain. 2010 Nov;11(11):1129-35.

Trial Protocol

18. Timmerman H, Wilder-Smith O, van Weel C, Wolff A, Vissers K. Detecting the neuropathic pain component in the clinical setting: a study protocol for validation of screening instruments for the presence of a neuropathic pain component. BMC Neurol. 2014 May 2;14:94. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4046010